

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 19:58:54 ; Search time 1988 Seconds  
(without alignments)  
244.512 Million cell updates/sec

Title: US-09-758-881-115

Perfect score: 20

Sequence: 1 gccacgacatctgctgcttc 20

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 33330

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: em\_estda:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estinu:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: qb\_est1:\*  
10: qb\_est2:\*  
11: qb\_hic:\*  
12: qb\_est3:\*  
13: qb\_est4:\*  
14: qb\_est5:\*  
15: em\_estlin:\*  
16: em\_estlom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inu:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_png:\*  
27: em\_gss\_vrt:\*  
28: qb\_gss1:\*  
29: qb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 13.4  | 67.0        | 29     | 28    | AZ780164 2M0017E17 |
| 2          | 12.2  | 61.0        | 25     | 9     | A1748295 sb50f02.Y |
| 3          | 12    | 60.0        | 24     | 28    | AZ779573 2M0016K09 |
| 4          | 12    | 60.0        | 27     | 28    | AZ404206 1M0172120 |

| 5  | 12   | 60.0 | 29 | 29 | TA141E08P | AL466622 T. brucei  |
|----|------|------|----|----|-----------|---------------------|
| 6  | 11.8 | 59.0 | 24 | 29 | B2358821  | B2358821 SALK 1333  |
| 7  | 11.6 | 58.0 | 22 | 28 | A2976330  | A2976330 2M0251P08  |
| 8  | 11.6 | 58.0 | 24 | 28 | A2826814  | A2826814 2M0102M15  |
| 9  | 11.6 | 58.0 | 28 | 9  | A1444428  | A1444428 fb38c12.x  |
| 10 | 11.6 | 58.0 | 28 | 9  | A1756191  | A1756191 FTESTga40  |
| 11 | 11.2 | 56.0 | 28 | 9  | A1755903  | A1755903 FTESTga08  |
| 12 | 11.2 | 55.0 | 22 | 28 | A2864977  | A2864977 2M0174D21  |
| 13 | 11.1 | 55.0 | 28 | 9  | AA717506  | AA717506 vu22e03.r  |
| 14 | 11.1 | 55.0 | 28 | 29 | DME546620 | AJ546620 Drosophila |
| 15 | 11.1 | 55.0 | 30 | 28 | A2396426  | A2396426 1M0160N09  |
| 16 | 11.1 | 55.0 | 30 | 28 | BH810436  | BH810436 SALK_0495  |
| 17 | 11.1 | 55.0 | 30 | 28 | BH847383  | BH847383 SALK_0531  |
| 18 | 10.8 | 54.0 | 19 | 28 | A2585898  | A2585898 1M0391122  |
| 19 | 10.8 | 54.0 | 24 | 28 | A2394196  | A2394196 1M0157H08  |
| 20 | 10.8 | 54.0 | 25 | 28 | A2782142  | A2782142 2M0022H10  |
| 21 | 10.8 | 54.0 | 29 | 28 | A2449708  | A2449708 1M0248M04  |
| 22 | 10.6 | 53.0 | 19 | 28 | A2430028  | A2430028 1M0214116  |
| 23 | 10.6 | 53.0 | 21 | 28 | A2786813  | A2786813 2M0032K11  |
| 24 | 10.6 | 53.0 | 26 | 28 | A2657494  | A2657494 1M0533A15  |
| 25 | 10.6 | 53.0 | 28 | 14 | T47368    | T47368 yD13B04.r1   |
| 26 | 10.6 | 53.0 | 29 | 12 | BM396481  | BM396481 5009-0-21  |
| 27 | 10.6 | 53.0 | 29 | 28 | A2854411  | A2854411 2M0158B05  |
| 28 | 10.4 | 52.0 | 19 | 28 | A2613058  | A2613058 1M041C18   |
| 29 | 10.4 | 52.0 | 20 | 28 | A2309949  | A2309949 1M0017M22  |
| 30 | 10.4 | 52.0 | 22 | 29 | PCB304213 | AJ304213 P15modiu   |
| 31 | 10.4 | 52.0 | 28 | 9  | A1756191  | A1756191 FTESTga40  |
| 32 | 10.4 | 52.0 | 29 | 28 | A2769892  | A2769892 1M0571C03  |
| 33 | 10.2 | 51.0 | 19 | 28 | A2858978  | A2858978 2M0164F24  |
| 34 | 10.2 | 51.0 | 25 | 9  | A1594892  | A1594892 vgl2c06.x  |
| 35 | 10.2 | 51.0 | 25 | 9  | A1743387  | A1743387 wq92f01.x  |
| 36 | 10.2 | 51.0 | 26 | 28 | A2620130  | A2620130 1M0452A14  |
| 37 | 10.2 | 51.0 | 28 | 12 | BM399938  | BM399938 5009-0-73  |
| 38 | 10.2 | 51.0 | 30 | 10 | BG424013  | BG424013 602447475  |
| 39 | 10.2 | 51.0 | 22 | 28 | A2492512  | A2492512 1M0326K08  |
| 40 | 10.0 | 50.0 | 24 | 28 | A2308017  | A2308017 1M0010M05  |
| 41 | 10.0 | 50.0 | 25 | 28 | A2364381  | A2364381 1M0110A08  |
| 42 | 10.0 | 50.0 | 26 | 28 | A2386258  | A2386258 1M0145E08  |
| 43 | 10.0 | 50.0 | 27 | 28 | A2797359  | A2797359 2M005306   |
| 44 | 10.0 | 50.0 | 27 | 28 | A2876196  | A2876196 2M0191A12  |
| 45 | 10.0 | 50.0 | 28 | 28 | A2596903  | A2596903 1M0410024  |

## ALIGNMENTS

RESULT 1  
LOCUS AZ780164 29 bp UNA linear GSS 16-FEB-2001  
DEFINITION 2M0017E17F Mouse 10kb plasmid UUCIM library Mus musculus genomic  
clone UUCG2M0017E17 F. genomic survey sequence.

ACCESSION AZ780164  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

GSS.  
Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts

JOURNAL  
COMMENT

Unpublished  
Contact: Robert B. Weiss  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT  
84112, USA  
Tel.: 801 585 5606  
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu  
 Insert length: 10000 Std Error: 0.00  
 Plate: 0017 Row: E Column: 17  
 Seq primer: CGTGTAAACGACGACGCT  
 Class: plasmid ends  
 High quality sequence stop: 29.

## FEATURES

## SOURCE

1..29  
 Location/Qualifiers

/organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="U00C2M0017E17"  
 /sex="Male"  
 /lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_1lb="Mouse 10kb plasmid U00C1M 1library"  
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (91:473214191AF12072.1), a copy number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## HASH COUNT

9 a 4 c 10 g 6 t

## ORIGIN

## Query Match

Best Local Similarity 93.3%; Pred. No. 1e+05; Length 29;

## Matches

14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTCACGATCTGCTG 16

Db 27 CTCACGATCTGCTG 13

## RESULT 2

## LOCUS

## DEFINITION

A1748295 25 bp mRNA linear EST 30-NOV-2001  
 sb50102.Y1 Gm-c1011 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 Gm-c1011-340 5' similar to SW:GIC1 SOYBN P04776 GLYCININ G1  
 PRECURSOR [CONTAINS: GLYCININ A1A SUBUNIT; GLYCININ BX SUBUNIT]. ;  
 mRNA sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

A1748295.1 GI:5126559  
 EST.  
 Glycine max (soybean)  
 Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids  
 ; eustrochylid; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine;  
 1 (bases 1 to 25)

## REFERENCE

## AUTHORS

Shoemaker, R., Keim, P., Vodkin, L., Erpelidg, J., Corvelli, V., Khanna  
 A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers  
 Y., Person, B., Waller, T., Gibbons, M., Page, D., Harvey, N., Schurk  
 R., Riller, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
 R., Waterston, R. and Wilson, R.  
 Public Soybean EST Project  
 Unpublished  
 Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: estewartson.wustl.edu

Trace considered overall poor quality  
 possible reversed clone; similarity on wrong strand This clone is  
 available through: Resgen, Invitrogen Corp. 2130 South Memorial  
 Parkway Huntsville, AL 35801 For further information call: (800  
 )-533-4363 or contact via email: ccuterresgen.com  
 Insert length: 2008 Std Error: 0.00  
 High quality sequence stop: 1.

## FEATURES

## SOURCE

1..25  
 Location/Qualifiers

/organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1011-340"  
 /tissue\_type="immature cotyledons of greenhouse grown  
 plants"  
 /lab\_host="DH10B"  
 /clone\_1lb="Gm-c1011"  
 /note="Vector: pBluescript II SK(+), Site-1: EcoRI; Site-2:  
 XhoI. This cDNA library was constructed from mRNA isolated  
 from immature cotyledons (100-200mg) of greenhouse grown  
 plants. The cDNA library was prepared using the lile  
 Technologies SuperScript cDNA library construction kit.  
 Complementary DNA was synthesized from mRNA using a poly  
 (dT) sequence with a Not I restriction site. Sal I  
 linkers adapters were ligated to the blunt-ended cDNA  
 fragments followed by Not I digestion. The cDNA fragments  
 were directionally cloned into the NotI-Sal I restriction  
 site of the pSPORT 1 vector. The ligated cDNA fragments  
 were transformed into E. coli Electromax DH10B host cells.  
 This library was constructed by Dr. Lila Vodkin and Dr.  
 Anu Khanna."

## HASH COUNT

2 a 5 c 4 g 14 t

## ORIGIN

## Query Match

Best Local Similarity 61.0%; Score 12.2; DB 9; Length 25;

## Matches

14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 TCCAGCATCTGCTGCT 19

Db 2 TTCAGTAGCTGCTGCT 18

## RESULT 3

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

A2779573 24 bp DNA linear GSS 16-FEB-2001  
 ZM0016K09F Mouse 10kb plasmid U00C1M library Mus musculus genomic  
 clone U00C2M0016K09 F, genomic survey sequence.  
 A2779573  
 A2779573.1 GI:12910362  
 GSS.  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 24)

## REFERENCE

## AUTHORS

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,  
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Kelly  
 M., Rose, R., Rose, R., Stokes, F., Tinney, A., von Niederhausern, A.  
 and Wright, D., Weiss, R.  
 Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts  
 Unpublished  
 Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg, 20 S 2030 E, SLC, UT

84112, USA  
 Tel.: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0016 row: K column: 09  
 Seq primer: CGTTGTAAAAACGACGGCTCAGT  
 Class: plasmid ends  
 High quality sequence stop: 24.

## FEATURES

## source

1. 24  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="U08C2M016K09"  
 /sex="Male"  
 /lab\_host="E. coli strain XL10-Gold, TI-resistant, F-"  
 /clone\_lib="Mouse 10kb plasmid U08C2M library"  
 /note="Vector: PMD42nv. Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (911473211419b/AP129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## BASE COUNT

8 a 8 c 8 g 0 t

## ORIGIN

Query Match 60.0%; Score 12; DB 28; Length 24;  
 Best Local Similarity 75.0%; Pred. No. 3.3e+05;  
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

## QY

1 GCTCCAGCAGTCGCTGCTTC 20

## DB

23 GCTGCTGCTGCTGCTGCTGC 4

RESULT 4  
 LOCUS A2404206 27 bp DNA linear GSS 03-OCT-2000  
 DEFINITION 1M0172120F Mouse 10kb plasmid U08C2M library Mus musculus genomic clone U08C2M0172120 F, genomic survey sequence.  
 ACCESSION A2404206  
 VERSION A2404206.1 GI:10528219  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 27)  
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R.

## AUTHORS

## REFERENCE

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

## TITLE

## JOURNAL

Unpublished  
 Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah

Rm. 308, Biomedical Polymers Research Bldg. 20 S 2030 E, SLG, UT  
 84112, USA  
 Tel.: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0172 row: I column: 20  
 Seq primer: CGTTGTAAAAACGACGGCTCAGT  
 Class: plasmid ends  
 High quality sequence stop: 27.

## FEATURES

## source

1. 27  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="U08C1M0172120"  
 /sex="Male"  
 /lab\_host="E. coli strain XL10-Gold, TI-resistant, F-"  
 /clone\_lib="Mouse 10kb plasmid U08C1M library"  
 /note="Vector: PMD42nv. Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (911473211419b/AP129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## BASE COUNT

0 a 9 c 9 g 9 t

## ORIGIN

Query Match 60.0%; Score 12; DB 28; Length 27;  
 Best Local Similarity 75.0%; Pred. No. 3.4e+05;  
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

## QY

1 GCTCCAGCAGTCGCTGCTTC 20

## DB

3 GCTGCTGCTGCTGCTGCTGC 22

RESULT 5  
 LOCUS TAI141E08P/c 29 bp DNA linear GSS 13-DEC-2000  
 DEFINITION T. brucei sheared genomic DNA clone 141e08, forward sequence.  
 ACCESSION AL466622  
 VERSION AL466622.1 GI:11835977  
 KEYWORDS GSS.  
 SOURCE Trypanosoma brucei  
 ORGANISM Trypanosoma brucei  
 Eukaryota; Euzlenzoa; Kinetoplastida; Trypanosomatidae;

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA. E-mail: barrrell@sanger.ac.uk and nh@sanger.ac.uk  
 Constructed at the Institute for Genomic Research (TIGR).



adaplored vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 11 a 2 c 6 g 3 t

Query Match 58.0%; Score 11.6; DB 28; Length 22;  
Best Local Similarity 77.8%; Pred. No. 4.6e+05;  
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0.

QY 2 CTCACGATCTGCTCTT 19  
11 11111111 1111  
Db 19 CTCACGATCTTCTT 2

RESULT 8 24 bp DNA linear GSS 20-FEB-2001  
LOCUS A2826814

DEFINITION 2M0102M15R Mouse 10kb plasmid mmgclm library Mus musculus genomic clone UUGC2M0102M15 R, genomic survey sequence.

ACCESSION A2826814  
VERSION A2826814.1 GI:12996722

KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE 1 (bases 1 to 24)  
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.

AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Kelly, M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausern, A. and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL. Unpublished  
COMMENT. Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., STC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: dunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0102 ROW: M Column: 15  
Seq primer: CACACAGCAACACGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 24.

FEATURES  
SOURCE Location/Qualifiers

1..24  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC2M0102M15"  
/sex="Male"  
/lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/note="Vector: pMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/naares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g1:4732114[gbl/AP129072.1], a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and

purified. The sheared, adaptored mouse DNA was annealed to adaplored vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 1 a 9 c 5 g 9 t

Query Match 58.0%; Score 11.6; DB 28; Length 24;  
Best Local Similarity 77.8%; Pred. No. 4.7e+05;  
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 TCCAGCATCTGCTCTT 20  
11111 11111111  
Db 1 TCCAGCCTCTGTGCTGC 18

RESULT 9  
LOCUS A1444428

DEFINITION A1444428 28 bp mRNA linear EST 07-JUN-2001  
FB38c12.X1 zebrafish Washu MP1MG EST Danio rerio cDNA clone IMAGE:3714166.3 similar to SW:RL5\_RAT F09895 60S RIBOSOMAL PROTEIN L5, mRNA sequence.

ACCESSION A1444428  
VERSION A1444428.1 GI:4281620

KEYWORDS EST.

SOURCE Danio rerio (zebrafish)

ORGANISM

REFERENCE 1 (bases 1 to 28)  
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi; Actinopterygii; Neopterygii; teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

AUTHORS Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McRann, P., Waterston, P. and Wilson, R.

TITLE Washu zebrafish EST Project 1998  
JOURNAL. Unpublished  
COMMENT. Contact: Stephen L. Johnson  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: zbrfish@watson.wustl.edu  
CNA Library Preparation: Matthew Clark, cDNA Library Arrayed by:  
Matthew Clark, DNA Sequencing by: Washington University Genome  
Sequencing Center Clone Distribution: Genome Systems, St. Louis,  
Missouri (web address: www.genomesystems.com) (email contact:  
info@genomesystems.com) and Research Genetics, Huntsville, Alabama  
(web address: www.resgen.com) (email contact: info@resgen.com) and  
Researchzentrum Primatendatenbank, Berlin, Germany (web address:  
www.rzp.dg.de)

Trace considered overall poor quality  
Possible reversed clone: similarity on wrong strand  
Seq primer: T7 ET from Amersham  
High quality sequence stop: 1  
POLYA-No.

FEATURES  
SOURCE Location/Qualifiers

1..28  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="IMAGE:3714166"  
/sex="mixed"  
/tissue\_type="26 somite embryos, adult livers, shield  
stage embryos"  
/lab\_host="XLI-blue MRF"  
/clone\_lib="zebrafish Washu MP1MG EST"  
/note="Vector: pSPOR1; site:1: NotI; site:2: SalI; 1st  
strand cDNA was primed with a Not I - oligo(drp15 primer  
[5'-pGCTAGTTCATGATGCGACGCGCGCTTTTCTTTTCTTTT3'];  
double-stranded cDNA was ligated to Sal I adaptors (BRL),

digested with Not I and cloned into the Not I and Sal I sites of the pSPORT1 vector (BRL). Library was constructed by Matthew Clark (Lehrach lab; ICRF, London and Max Planck Institut fuer Molekulare Genetik, Berlin). cDNAs for EST analysis were selected following oligonucleotide hybridization fingerprinting of arrayed clones from zebrafish late somitogenesis (26 ss), adult liver or embryonic shield stage (5.6 h) libraries. Fingerprint data were used to computationally cluster cDNAs, and a single cDNA from each cluster was chosen for sequencing. In some cases multiple members of the same cluster were sequenced to assess clustering parameters or single clones were sequenced additional times to assess quality control."

BASE COUNT  
ORIGIN  
Query Match 58.0%; Score 11.6; DB 9; Length 28;  
Best Local Similarity 77.8%; Pred. No. 5e+05;  
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CTCGACATCTGCTGCT 19  
||||| 1 ||||| 11  
Db 10 CTCGACATCTGCTGCT 27

RESULT 10  
A1756191 28 bp mRNA linear EST 18-JAN-2000  
LOCUS E18T8a40h07.y1 Eimeria S5-2 Sporozoite stage Eimeria tenella cDNA  
DEFINITION 5' similar to U18-064507 064507 SERINE 1 ULTRA HIGH SULFUR PROTEIN.  
; mRNA sequence.

ACCESSION A1756191  
VERSION A1756191.1 GI:5149914  
KEYWORDS EST.  
SOURCE Eimeria tenella  
ORGANISM Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;

REFERENCE 1 (bases 1 to 28)  
AUTHORS

LIBERATOR, P., DIAZ, C., TANG, K., MARRA, M., HILLIER, L., KUCABA, T.,  
MARTIN, J., WYLLIE, T., UNDERWOOD, K., STEPTOE, M., THEISING, B., ALLEN,  
M., BOWERS, Y., PERSON, B., SWALLER, T., GIBBONS, M., PAPE, D., HARVEY,  
N., SCHURK, R., RITTER, E., KOHN, S., FLORENCE, N., SHIN, T., JACKSON,  
Y., CARDENAS, M., MCCANN, R., WATERSTON, R., WILSON, R. and SIBLEY, D.  
WashU-Merck Eimeria tenella project

Unpublished  
Contact: David Sibley, Ph.D.  
WashU-Merck Eimeria tenella project

Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@watson.wustl.edu

Contact David Sibley (toxoc@est@watson.wustl.edu) for further  
information relating to organism, libraries, or clone availability.  
Trace considered overall poor quality  
Possible reversed clone: similarity on wrong strand  
Seq primer: -40RP from Glbco  
High quality sequence stop: 1.

#### FEATURES

Location/Qualifiers  
1..28  
/organism="Eimeria tenella"  
/mol\_type="mRNA"  
/strain="LS18"  
/db\_xref="taxon:5802"  
/dev\_stage="Sporozoite"  
/lab\_host="SOLR R. coli"  
/clone\_lib="Eimeria S5-2 Sporozoite stage"  
/note="Vector: Bluescript SK-; Site 1: EcoRI; Site 2: XhoI  
; Sporozoites were obtained from in vitro sporulated  
excysted oocysts of E. tenella grown in chickens. cDNA  
was synthesized from poly mRNA using an oligo-dT primer

containing a XhoI site. Following second strand synthesis,  
PfuRI adapters were ligated to the cDNA and products were  
size-selected on Sephadryl 500. cDNAs were digested with  
EcoRI/XhoI and cloned into Lambda Zap II (Stratagene).  
Clones were converted to phagemids by mass excision using  
Exassist helper phage and SOLR cells (Stratagene).  
Insert sizes range from 1.2-2.9 Kb."

BASE COUNT  
ORIGIN  
Query Match 58.0%; Score 11.6; DB 9; Length 28;  
Best Local Similarity 77.8%; Pred. No. 5e+05;  
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCTCCGACATCTGCTGCT 18  
||||| 1 ||||| 11  
Db 18 GCTCCGACATCTGCTGCT 1

RESULT 11  
A1755903 28 bp mRNA linear EST 18-JAN-2000  
LOCUS E18T8a08h09.y1 Eimeria M5-6 Merzoite stage Eimeria tenella cDNA  
DEFINITION 5' similar to W1-TU1H7.8 CHU3532,, mRNA sequence.

ACCESSION A1755903  
VERSION A1755903.1 GI:5149626  
KEYWORDS EST.  
SOURCE Eimeria tenella  
ORGANISM Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;

REFERENCE 1 (bases 1 to 28)  
AUTHORS

LIBERATOR, P., DIAZ, C., TANG, K., MARRA, M., HILLIER, L., KUCABA, T.,  
MARTIN, J., WYLLIE, T., UNDERWOOD, K., STEPTOE, M., THEISING, B., ALLEN,  
M., BOWERS, Y., PERSON, B., SWALLER, T., GIBBONS, M., PAPE, D., HARVEY,  
N., SCHURK, R., RITTER, E., KOHN, S., FLORENCE, N., SHIN, T., JACKSON,  
Y., CARDENAS, M., MCCANN, R., WATERSTON, R., WILSON, R. and SIBLEY, D.  
WashU-Merck Eimeria tenella project

Unpublished  
Contact: David Sibley, Ph.D.  
WashU-Merck Eimeria tenella project

Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@watson.wustl.edu

Contact David Sibley (toxoc@est@watson.wustl.edu) for further  
information relating to organism, libraries, or clone availability.  
Trace considered overall poor quality  
Possible reversed clone: similarity on wrong strand  
Seq primer: -40RP from Glbco  
High quality sequence stop: 1.

#### FEATURES

Location/Qualifiers  
1..28  
/organism="Eimeria tenella"  
/mol\_type="mRNA"  
/strain="LS18"  
/db\_xref="taxon:5802"  
/dev\_stage="Merzoite"  
/lab\_host="SOLR R. coli"  
/clone\_lib="Eimeria M5-6 Merzoite stage"  
/note="Vector: Bluescript SK-; Site 1: EcoRI; Site 2: XhoI  
; Merzoites were obtained from ceecal scrapings of  
chickens infected with E. tenella. The library may  
contain a small percentage of host or bacterial  
contaminants. cDNA was synthesized from poly mRNA using  
an oligo-dT primer containing a XhoI site. Following  
second strand synthesis, EcoRI adapters were ligated to  
the cDNA and products were size-selected on Sephadryl  
500. cDNAs were digested with EcoRI/XhoI and cloned into  
Lambda Zap II (Stratagene). Clones were converted to  
phagemids by mass excision using Exassist helper phage and  
SOLR cells (Stratagene). Insert sizes range from 0.7-1.5

|            |  |   |  |               |           |   |
|------------|--|---|--|---------------|-----------|---|
|            | BASE COUNT   | 3 a   | 7 c  | 5 g           | 7 t       | Chemically-competent E. coli XL10-GOLD (Stratagene) cells and selected for ampicillin resistance. |
| OY         | Query Match<br>Best Local Similarity   | 8 CATCTGCTGCT 18<br>     <br>11: Conservative | 55.0%; Score 11; DB 28; Length 22;<br>100.0%; Pred. No. 7.7e+05; | 0; Mismatches | 0; Indels | 0; Gaps   |
| Db         | 2 CATCTGCTGCT 12   |   |  |               |           |   |
| RESULT 13  |  |   |  |               |           |   |
| LOCUS      | AA717506   |   |  |               |           |   |
| DEFINITION | AA717506 28 bp mRNA linear EST 29-DEC-1997<br>vuz2e03.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone IMAGE:118104 5' similar to TR:060961 Q60961 GOLGI 4-TRANSMEMBRANE SPANNING TRANSPORTER MTP. ;, mRNA sequence.  |   |  |               |           |   |
| ACCESSION  | AA717506   |   |  |               |           |   |
| VERSION    | EST.   |   |  |               |           |   |
| KEYWORDS   | AA717506.1 GI:2729780  |   |  |               |           |   |
| SOURCE     | Mus musculus (house mouse)   |   |  |               |           |   |
| ORGANISM   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scurionathi; Muridae; Murinae; Mus.  |   |  |               |           |   |
| REFERENCE  | 1 (bases 1 to 28)<br>Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Thelsting,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.<br>The WashU-HMI Mouse EST Project<br>Unpublished  |   |  |               |           |   |
| AUTHORS    | Contact: Marra M/Mouse EST Project<br>WashU-HMI Mouse EST Project<br>Washington University School of Medicine<br>4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108<br>Tel: 314 286 1800<br>Fax: 314 286 1810<br>Email: mouseest@watson.wustl.edu   |   |  |               |           |   |
| TITLE      | This clone is available royalty-free through LNL ; contact the IMAGE Consortium (InfoImage.lnl.gov) for further information.<br>MGI:639252   |   |  |               |           |   |
| JOURNAL    | trace considered overall poor quality<br>possible reversed clone; similarity on wrong strand<br>Seq primer: -28ml3 rev2 ET from Amersham<br>High quality sequence stop: 1.   |   |  |               |           |   |
| COMMENT    | Location/Qualifiers<br>1..28<br>/organism="Mus musculus"<br>/mol_type="mRNA"<br>/strain="C3H"<br>/db_xref="taxon:10090"<br>/clone="IMAGE:118104"<br>/cell_line="C2C12"<br>/lab_host="DH10B"<br>/clone_lib="Barstead mouse myotubes MPLRB5"<br>/note="Vector: pF7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'<br>TCGTACGAATCGAAGTCGCAGCGCCCCCTTTTTTTTTTTTTTTTTT<br>3']; double-stranded cDNA was ligated to Eco RI adaptors [AAATGGAGCTCTTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pF7T3 vector.<br>Library constructed by Bob Barstead. The C2C12 cell line (available from ATCC, catalog # CRL-1772) differentiates rapidly, forming contractile myotubes and producing characteristic muscle proteins." |   |  |               |           |   |
| FEATURES   |  |   |  |               |           |   |
| SOURCE     |  |   |  |               |           |   |
| BASE COUNT | 4 a  | 11 c  | 8 g  | 5 t           |           |   |
| ORIGIN     |  |   |  |               |           |   |

|             |                  |                    |       |               |
|-------------|------------------|--------------------|-------|---------------|
| Query Match | 55.0%            | Score 11;          | DB 9; | Length 28;    |
| Best local  | Similarity 73.7% | Pred. No. 8.4e+05; |       |               |
| Matches     | 14;              | Conservative       | 0;    | Mismatches 5; |
|             |                  |                    |       | Indels 0;     |
|             |                  |                    |       | Gaps 0        |
| QY          | 2                | CTCCACACATCTGCTGTC | 20    |               |
|             |                  |                    |       |               |
| Db          | 10               | CTACACACCCGCTGCTGC | 28    |               |

|            |   |
|------------|---|
| RESULT 14  |   |
| DME546620  |   |
| LOCUS      | DME546620 28 bp DNA linear GSS 24-PEB-2003                              |
| DEFINITION | Drosophila melanogaster flanking sequence of RS element insertion site. |
| ACCESSION  | AF546620  |
| VERSION    | AF546620.1 GI:28554721  |
| KEYWORDS   | GSS: genome survey sequence.  |
| SOURCE     | Drosophila melanogaster (fruit fly)                                     |
| ORGANISM   | Drosophila melanogaster   |

REFERENCE  
AUTHORS

|            |   |
|------------|---|
| TITLE:     | Mapping of RS P element insertions in <i>Drosophila melanogaster</i> for the Drosdel second generation deficiency kit |
| JOURNAL:   | Unpublished   |
| REFERENCE: | 2 (bases 1 to 28)   |
| AUTHORS:   | Buckner J   |

|                    |  |
|--------------------|--|
| JOURNAL, REFERENCE | Unpublished  |
| AUTHORS            | 2 (bases 1 to 28)  |
| TITLE              | Ryder E.J.   |
| JOURNAL            | Direct Submission  |
| COMMENT            | Submitted (17-FEB-2003) Ryder E.J., Department of Genetics, University of Cambridge, Downing Street, CB2 3EH, UNITED KINGDOM<br>The insertion point of the p element is before base 1 of the |

| FEATURES            | COMMENT  |
|---------------------|--|
| location/Qualifiers | The insertion point of the p element is before base 1 of the sequence. Further information about this p element insertion line can be found at <a href="http://www.flyseq.org.uk">http://www.flyseq.org.uk</a> and <a href="http://www.drosdel.org.uk">http://www.drosdel.org.uk</a> . |

```

/misc-feature
1..28
/organism="Drosophila melanogaster"
/mol-type="genomic DNA"
/db_xref="taxon:7227"
/chromosome="4R"
/clone="p(RS)5-S2-3275"
/clone_1b="p(RS5)"
/note="read-5' end"

```

| BASE COUNT<br>ORIGIN | 1 a | 12 c | 5 g | 10 t | /note- <sup>a</sup> element insertion in the 3' to 5' orientation <sup>a</sup> |
|----------------------|-----|------|-----|------|--|
| 1                    | 1   | 1    | 1   | 1    | 1  |
| 2                    | 1   | 1    | 1   | 1    | 1  |
| 3                    | 1   | 1    | 1   | 1    | 1  |
| 4                    | 1   | 1    | 1   | 1    | 1  |
| 5                    | 1   | 1    | 1   | 1    | 1  |
| 6                    | 1   | 1    | 1   | 1    | 1  |
| 7                    | 1   | 1    | 1   | 1    | 1  |
| 8                    | 1   | 1    | 1   | 1    | 1  |
| 9                    | 1   | 1    | 1   | 1    | 1  |
| 10                   | 1   | 1    | 1   | 1    | 1  |
| 11                   | 1   | 1    | 1   | 1    | 1  |
| 12                   | 1   | 1    | 1   | 1    | 1  |
| 13                   | 1   | 1    | 1   | 1    | 1  |
| 14                   | 1   | 1    | 1   | 1    | 1  |
| 15                   | 1   | 1    | 1   | 1    | 1  |
| 16                   | 1   | 1    | 1   | 1    | 1  |
| 17                   | 1   | 1    | 1   | 1    | 1  |
| 18                   | 1   | 1    | 1   | 1    | 1  |
| 19                   | 1   | 1    | 1   | 1    | 1  |
| 20                   | 1   | 1    | 1   | 1    | 1  |
| 21                   | 1   | 1    | 1   | 1    | 1  |
| 22                   | 1   | 1    | 1   | 1    | 1  |
| 23                   | 1   | 1    | 1   | 1    | 1  |
| 24                   | 1   | 1    | 1   | 1    | 1  |
| 25                   | 1   | 1    | 1   | 1    | 1  |
| 26                   | 1   | 1    | 1   | 1    | 1  |
| 27                   | 1   | 1    | 1   | 1    | 1  |
| 28                   | 1   | 1    | 1   | 1    | 1  |
| 29                   | 1   | 1    | 1   | 1    | 1  |
| 30                   | 1   | 1    | 1   | 1    | 1  |
| 31                   | 1   | 1    | 1   | 1    | 1  |
| 32                   | 1   | 1    | 1   | 1    | 1  |
| 33                   | 1   | 1    | 1   | 1    | 1  |
| 34                   | 1   | 1    | 1   | 1    | 1  |
| 35                   | 1   | 1    | 1   | 1    | 1  |
| 36                   | 1   | 1    | 1   | 1    | 1  |
| 37                   | 1   | 1    | 1   | 1    | 1  |
| 38                   | 1   | 1    | 1   | 1    | 1  |
| 39                   | 1   | 1    | 1   | 1    | 1  |
| 40                   | 1   | 1    | 1   | 1    | 1  |
| 41                   | 1   | 1    | 1   | 1    | 1  |
| 42                   | 1   | 1    | 1   | 1    | 1  |
| 43                   | 1   | 1    | 1   | 1    | 1  |
| 44                   | 1   | 1    | 1   | 1    | 1  |
| 45                   | 1   | 1    | 1   | 1    | 1  |
| 46                   | 1   | 1    | 1   | 1    | 1  |
| 47                   | 1   | 1    | 1   | 1    | 1  |
| 48                   | 1   | 1    | 1   | 1    | 1  |
| 49                   | 1   | 1    | 1   | 1    | 1  |
| 50                   | 1   | 1    | 1   | 1    | 1  |
| 51                   | 1   | 1    | 1   | 1    | 1  |
| 52                   | 1   | 1    | 1   | 1    | 1  |
| 53                   | 1   | 1    | 1   | 1    | 1  |
| 54                   | 1   | 1    | 1   | 1    | 1  |
| 55                   | 1   | 1    | 1   | 1    | 1  |
| 56                   | 1   | 1    | 1   | 1    | 1  |
| 57                   | 1   | 1    | 1   | 1    | 1  |
| 58                   | 1   | 1    | 1   | 1    | 1  |
| 59                   | 1   | 1    | 1   | 1    | 1  |
| 60                   | 1   | 1    | 1   | 1    | 1  |
| 61                   | 1   | 1    | 1   | 1    | 1  |
| 62                   | 1   | 1    | 1   | 1    | 1  |
| 63                   | 1   | 1    | 1   | 1    | 1  |
| 64                   | 1   | 1    | 1   | 1    | 1  |
| 65                   | 1   | 1    | 1   | 1    | 1  |
| 66                   | 1   | 1    | 1   | 1    | 1  |
| 67                   | 1   | 1    | 1   | 1    | 1  |
| 68                   | 1   | 1    | 1   | 1    | 1  |
| 69                   | 1   | 1    | 1   | 1    | 1  |
| 70                   | 1   | 1    | 1   | 1    | 1  |
| 71                   | 1   | 1    | 1   | 1    | 1  |
| 72                   | 1   | 1    | 1   | 1    | 1  |
| 73                   | 1   | 1    | 1   | 1    | 1  |
| 74                   | 1   | 1    | 1   | 1    | 1  |
| 75                   | 1   | 1    | 1   | 1    | 1  |
| 76                   | 1   | 1    | 1   | 1    | 1  |
| 77                   | 1   | 1    | 1   | 1    | 1  |
| 78                   | 1   | 1    | 1   | 1    | 1  |
| 79                   | 1   | 1    | 1   | 1    | 1  |
| 80                   | 1   | 1    | 1   | 1    | 1  |
|                      |     |      |     |      |  |

|                       |                 |                    |           |            |
|-----------------------|-----------------|--------------------|-----------|------------|
| Query Match           | 55.08;          | Score 11;          | DB 29;    | length 28; |
| Best Local Similarity | 73.78;          | Pred. No. 8.4e+05; |           |            |
| Matches 14;           | Conservative 0; | Mismatches 5;      | Indels 0; | Gaps 0;    |

|     |   |                      |    |
|-----|---|----------------------|----|
| cy  | 2 | CTCCAGCAATCTGCTGCTTC | 20 |
|     |   |                      |    |
| 10b | 1 | CTCTCCCGGACTGCTGCTTC | 19 |

RESULT 15

A7.396226

### DEFINITION

11

ACCESSION AZ396226  
 DEPOSITION 10/11/2000

## KEYWORDS

**SOURCE:** *Mus musculus* (house mouse)

|           |  |
|-----------|--|
| ORGANISM  | Mus musculus   |
|           | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;    |
|           | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus    |
| REFERENCE | 1 (pages 1 to 30)  |
| AUTHORS   | Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., |

|         |  |
|---------|--|
| TITLE   | Mouse whole genome scaffolding with paired end reads from 10kb |
| JOURNAL | plasmid inserts  |
| COMMENT | Unpublished  |
|         | Contact: Robert B. Weiss                                       |

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert length: 1000 Std error: 0.00  
Plate: 0160 row: N column: 09  
Seq primer: CACACAGGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 30.

| FEATURES | location/Qualifiers |
|----------|---------------------|
| source   | 1. .30              |

BASE COUNT  
ORIGIN

6 a 7 c 5 g 12 t

```

/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGGCM160N09"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone.lib="Mouse 10kb plasmid UGGCM library"
/notes="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (gII47311419b)AFL13072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
```

```

Query Match:          55.0%; Score 11; DB 28; Length 30;
Best Local Similarity 100.0%; Pred. No. 8, 6e+05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 ACCATCTGCTG 16
    | | | | | | | | | |
    | | | | | | | | | |
Db 1 ACCATCTGCTG 11

```

```
Search completed: August 19, 2003, 21:20:41
Job time : 1993 secs
```